

OCA PAD INITIATION - PROJECT HEADER INFORMATION

04/27/88

Active

Project #: E-24-620
Center #: R5834-1A0

Cost share #:
Center shr #:

Rev #: 0
OCA file #:
Work type : RES
Document : GRANT
Contract entity: GTRC

Contract#: AFOSR-88-0137
Prime #:

Mod #:

Subprojects ? : N
Main project #:

Project unit: ISYE
Project director(s):
SERFOZO R F ISYE

Unit code: 02.010.124

Sponsor/division names: AIR FORCE
Sponsor/division codes: 104

/ BOLLING AFB, DC
/ 001

Award period: 880215 to 890214 (performance) 890414 (reports)

Sponsor amount	New this change	Total to date
Contract value	50,870.00	50,870.00
Funded	50,870.00	50,870.00
Cost sharing amount		0.00

Does subcontracting plan apply?: N

Title: STOCHASTIC FLOWS IN NETWORKS

PROJECT ADMINISTRATION DATA

OCA contact: Steven K. Watt 894-4820

Sponsor technical contact Sponsor issuing office

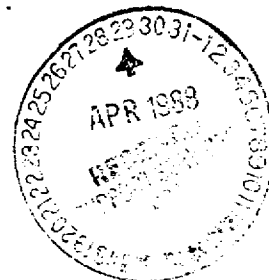
BRIAN W. WOODRUFF, MAJOR, USAF/NM
(202)767-5025
AFOSR
BUILDING 410
BOLLING AFB, D.C. 20332

TERRY D. MOEN/PKZ
(202)767-4957
AFOSR
BUILDING 410
BOLLING AFB, D.C. 20332

Security class (U,C,S,TS) : U
Defense priority rating : N/A
Equipment title vests with: Sponsor
NONE PROPOSED.

ONR resident rep. is ACO (Y/N): N
N/A supplemental sheet
GIT X

Administrative comments -
INITIATION. CONTINUATION OF E-24-659



NOTICE OF PROJECT CLOSEOUT

Project No. E-24-620

Center No. R5834-1A0

Project Director R. F. Serfozo

School/Lab ISyE

Sponsor Bolling Air Force Base

Contract/Grant No. AFOSR-88-0137

GTRC XX GIT

Prime Contract No. N/A

Title Stochastic Flows in Networks

Effective Completion Date 4/14/89 **(Performance)** 6/14/89 **(Repo**

None
☒ Final Invoice or Copy of Last Invoice
☒ Final Report of Inventions and/or Subcontracts- Patent questionnaire sent
☒ Government Property Inventory & Related Certificate
☐ Classified Material Certificate
☒ Release and Assignment
☐ Other

Includes Subproject No(s).

Subproject Under Main Project No.

Continues Project No. E-24-659

Continued by Project No.

<u>X</u>	Project Director
<u>X</u>	Administrative Network
<u>X</u>	Accounting
<u>X</u>	Procurement/GTRI Supply Services
<u>X</u>	Research Property Management
<u> </u>	Research Security Services

<u>X</u>	Reports Coordinator (OCA)
<u>X</u>	GTRC
<u>X</u>	Project File
<u>2</u>	Contract Support Division (OCA)
	Other

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION N/A			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY N/A			3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.		
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE N/A					
4. PERFORMING ORGANIZATION REPORT NUMBER(S)			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
6a. NAME OF PERFORMING ORGANIZATION Georgia Institute of Technology		6b. OFFICE SYMBOL (if applicable)		7a. NAME OF MONITORING ORGANIZATION AFOSR/NM	
6c. ADDRESS (City, State, and ZIP Code) School of Industrial and Systems Engineering Atlanta, GA 30332-0205			7b. ADDRESS (City, State, and ZIP Code) AFOSR/NM Bldg 410 Bolling AFB DC 20332-6448		
8a. NAME OF FUNDING / SPONSORING ORGANIZATION AFOSR		8b. OFFICE SYMBOL (if applicable) NM		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
8c. ADDRESS (City, State, and ZIP Code) AFOSR/NM Bldg 410 Bolling AFB DC 20332-6448			10. SOURCE OF FUNDING NUMBERS		
			PROGRAM ELEMENT NO. 61102F	PROJECT NO. 2304	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) Markovian Network Processes with System-Dependent Transition Rates					
12. PERSONAL AUTHOR(S) Richard F. Serfozo					
13a. TYPE OF REPORT Preprint		13b. TIME COVERED FROM 2/1/88 TO 11/10/88		14. DATE OF REPORT (Year, Month, Day) Nov. 10, 1988	
15. PAGE COUNT 37					
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) A Markovian network process describes the movement of discrete units among a set of nodes that process the units. There is considerable knowledge of such networks, often called queuing networks, in which the nodes operate independently and the routes of the units are independent. The focus of this study, in contrast, is on networks with dependent nodes and routings. Examples of dependencies are parallel processing across several nodes, alternate routing of units to avoid congestion, and accelerating or decelerating the processing rate at a node depending on downstream congestion. We introduce a rather general canonical network process and derive its equilibrium distribution. This distribution takes the form of an interchangeable product of functions of decreasing vectors. This new type of distribution is rather universal and may apply to other multi-variate processes as well. A basic idea in our approach is that we link certain micro-level balance properties of the network routing to the processing rates at the nodes. The link is via routing-balance partitions of nodes that are inherent in any network. We also give necessary and sufficient conditions under which a unit					
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS				21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL Maj. Brian Woodruff			22b. TELEPHONE (Include Area Code)		22c. OFFICE SYMBOL NM

Markovian Network Processes with System-Dependent Transition Rates

Richard F. Serfozo

Georgia Institute of Technology

Abstract

A Markovian network process describes the movement of discrete units among a set of nodes that process the units. There is considerable knowledge of such networks, often called queueing networks, in which the nodes operate independently and the routes of the units are independent. The focus of this study, in contrast, is on networks with dependent nodes and routings. Examples of dependencies are parallel processing across several nodes, blocking of transitions because of capacity constraints on nodes, alternate routing of units to avoid congestion, and accelerating or decelerating the processing rate at a node depending on downstream congestion. We introduce a rather general canonical network process and derive its equilibrium distribution. This distribution takes the form of an interchangeable product of functions of decreasing vectors. This new type of distribution is rather universal and may apply to other multi-variate processes as well. A basic idea in our approach is that we link certain micro-level balance properties of the network routing to the processing rates at the nodes. The link is via routing-balance partitions of nodes that are inherent in any network. We also give necessary and sufficient conditions under which a unit moving in the network sees a time average for the unmoved units. Finally, we discuss when certain flows between nodes in an open network are Poisson processes.

Keywords: Markovian network process, Queueing network, Dependent nodes and routing, Palm probability, Poisson process, Moving units see time averages.

1 Introduction

A stochastic network process is a model for a system in which discrete units move among a set of nodes that process the units. Such processes are often called queueing networks. Some archetypal stochastic networks are as follows:

Flexible Manufacturing Networks: Parts, tools or material (units) move among a group of work stations and storage areas (nodes) that machine and store the units for later use or for shipping.

Telecommunications Networks: Telephone calls, data packets or messages move among operators or switching stations.

Computer Networks: Transactions, data packets or programs move among processors, computers, peripheral equipment or files.

Maintenance Networks: Repairable parts or equipment needed for the operation of a large system move among locations where they are used, repaired, and stored.

Distribution Networks: Goods, orders or trucks move among plants, warehouses or market locations.

Biological Networks: Animals, cells, molecules, neurons, etc. move among locations, states or shapes.

Typical concerns associated with stochastic networks are: assessing the operational quality or feasibility of a network, comparing the quality of several prototype networks, designing a least-cost network (determining numbers of machines, tools or routes), and selecting optimal operating rules for the network. To address these issues requires an understanding of the probabilistic behavior of the network in terms of the equilibrium or stationary probability distribution of the numbers of units at the nodes. This distribution is used to derive various performance measures of the network such as the expected cost of operating the network or the percentage of time a sector of the network is overloaded. It is also a basic ingredient for the development of mathematical programming algorithms to select optimal network designs and operating rules. Other important network features include the rate of flow of units on the arcs and through the nodes (the throughputs), the time a unit spends in the network, and the time it takes for a unit to move from one sector of the network to another.

Most of the theory of Markovian network processes is for Jackson Network processes; the early papers are [2], [11], [14], and [28], and standard

textbook references are [10], [16], [26] and [29]. The key features of Jackson processes are:

- The units move one at a time.
- The nodes operate independently.
- The transition rates depend on only local information: the service rate at a node depends on only the number of units at that node and is independent of the rest of the network.
- The routes of units are independent of each other and hence independent of the congestion in the network.
- The equilibrium distribution is of product form.

Jackson processes can be viewed as the first generation of stochastic network processes. They play an important role in analyzing networks with independent nodes and routes.

By its very nature, however, a network is a system of interacting nodes in which the operation of a node and the routing of a unit may depend on what is happening throughout the network. Examples of dependencies are:

- Parallel processing across several nodes.
- Alternate routing of units to avoid congestion.
- Accelerating or decelerating the processing rate at a node whenever downstream nodes are starved or congested.
- Units are blocked from entering a sector of the network when the sector cannot handle any more units.

Such dependencies are omnipresent in the networks mentioned above. To model networks with dependencies will require a new generation of network processes that will typically have more complex, non-product-form equilibrium distributions. Some initial work on this theme has already begun [5], [6], [7], [12], [13], [16], [17], [18], [24], [26], [29]. We will comment on these references as we proceed.

The aim of the present paper is to introduce a fairly wide class of Markovian network processes with system-dependent transition rates that represent a variety of interactions between nodes. This class contains the Jackson processes and most of its generalizations developed to date. The equilibrium distributions for the processes take the form of an interchangeable product of functions of decreasing vectors. This new type of distribution appears to be rather universal and may apply to other processes as well. A key idea in our approach is that we link certain micro partial balance properties of the routing rates of the units to their processing rates at the nodes.

Our discussion proceeds as follows. In Sections 2 and 3, we present our notation and give preliminary examples and results. The point of Section 3 is to review two basic network processes that represent the two extreme ends of the spectrum of networks that is our focus. In Section 4 we define our network processes and derive their equilibrium distributions. We establish further properties of the processes in Sections 5, 6, 7, which cover the following topics:

- Blocking of certain transitions due to network constraints.
- Palm probabilities of a network at its transitions: When does a moving unit see a time average for the unmoved units?
- Poisson flows in open networks.

2 Notation

We shall consider a network consisting of m nodes, labeled $1, \dots, m$, that process discrete units that move among the nodes. We will confine our discussion to indistinguishable units and point out later how the results extend to multiple types of units. The evolution of the network is represented by the stochastic process $X = \{X_t : t \geq 0\}$ that records the numbers of units at the respective nodes. The state space of X is a set S of non-negative integer-valued vectors $n = (n_1, \dots, n_m)$, where n_j denotes the number of units at node j . The network is *open* if units enter it from outside and eventually exit; here S is the set of all non-negative vectors. The network is *closed* if there are a fixed number of units, say N , continually circulating in

it; here S is the set S_N of all $n = (n_1, \dots, n_m)$ such that $n_1 + \dots + n_m = N$. For simplicity, we assume throughout Sections 2-6 that X is such a closed network; Section 7 covers open networks. The results herein also have analogues for mixed networks (a combination of open and closed ones) and for network processes with more general state spaces as in (29) that monitor additional information about the nodes and units or environmental factors.

We shall assume that the network process X is a pure jump Markov process. The distribution of such a process is determined by its transition rates

$$q(n, n') = \lim_{t \downarrow 0} t^{-1} P\{X_t = n' \mid X_0 = n\}, \quad n' \neq n.$$

Our focus will be on networks in which only one unit may move at a time. Accordingly, we assume that a transition of X is synonymous with the movement of a single unit from one node to another. Specifically, if X is in state n and a unit moves from node j to node k , then the next state is $T_{jk}n$, where

$$T_{jk}n = \begin{cases} n - e_j + e_k & \text{for } n_j \geq 1 \\ n & \text{for } n_j = 0, \end{cases}$$

and e_j is the m -dimensional vector with 1 in entry j and 0's elsewhere. The distribution of the process is therefore determined by the transition rates $q(n, T_{jk}n)$, $j \neq k$, $n_j \geq 1$ ($q(n, n') = 0$ for all other states n'). We shall express these rates as

$$q(n, T_{jk}n) = \lambda_{jk} \phi_j(n) \Psi_{jk}(n), \quad (1)$$

where $\lambda_{jk} \geq 0$ and ϕ_j, Ψ_{jk} are positive functions on S . Section 5 covers the possibility of allowing $\phi_j(n)$ or $\Psi_{jk}(n)$ to be zero on part of S , which results in blocking.

The three-part factorization (1) is convenient, as we shall soon see, for isolating certain basic dependencies in the transitions. The particular factors and their interpretations will depend on the particular network. In general, one can view λ_{jk} as the *routing intensity* from j to k that is inherent in the structure of the network; it is typically a function of the node or arc characteristics, independent of the system state n . The λ_{jk} is sometimes the probability that a unit departing from node j is assigned to move to node k ; in this case $\lambda_{j1} + \dots + \lambda_{jm} = 1$ (we do not make this

restriction). The $\phi_j(n)$ can be viewed as the *departure rate* at node j . This may be the same as the service rate or the service rate multiplied by an expediting function as in Example 5.6. Typically $\phi_j(n)$ is a function of only a few n_k 's that have some relation to node j . Finally, we view $\Psi_{jk}(n)$ as a *system-dependent transfer rate* for arc j, k . It might be the intensity at which node k "pulls" or "accepts" units from j , or the intensity at which j "pushes" units to k . At this point, the intensities $\phi_j(n)$ and $\Psi_{jk}(n)$ are arbitrary functions. We will place restrictions on them in the course of our discussion.

Our only other general assumption on the λ_{jk} is that they be such that the network process X is irreducible. There is no loss in generality from this assumption. Since ϕ_j and Ψ_{jk} are positive, it is easy to see that X is irreducible if and only if the Markov matrix $\lambda_{jk} / \sum_{\ell=1}^m \lambda_{j\ell}$, $j, k = 1, \dots, m$, is irreducible. Consequently, there are positive numbers w_1, \dots, w_m that satisfy the *routing balance equations*

$$\sum_k (w_j \lambda_{jk} - w_k \lambda_{kj}) = 0. \quad (2)$$

The vector w_1, \dots, w_m is unique up to a constant multiple.

The assumption that the closed network process X is irreducible and the finiteness of its state space $S = S_N$ ensure that X has an equilibrium distribution, which we denote by $\pi(n)$, $n \in S$. That is, π is the unique probability measure that satisfies the total balance equations

$$\pi(n) \sum_{j,k} q(n, T_{jk}n) = \sum_{j,k} \pi(T_{jk}n) q(T_{jk}n, n), \quad n \in S. \quad (3)$$

We will often write this π as

$$\pi(n) = c \Phi(n) \prod_{j=1}^m w_j^{n_j}, \quad n \in S, \quad (4)$$

where w_1, \dots, w_m is as above, Φ is a positive function on S and c is a normalizing constant such that these $\pi(n)$'s sum to 1. Note that (4) is not a special form for π since any distribution can be written this way – simply set $\Phi(n) = \pi(n) / \prod_j w_j^{n_j}$ and $c = 1$.

In summary of the preceding discussion, $X = \{X_t : t \geq 0\}$ represents a closed, irreducible Markovian network process in which N units circulate

among m nodes, and the transition rates of X are

$$q(n, T_{jk}n) = \lambda_{jk} \phi_j(n) \Psi_{jk}(n).$$

The problem we address is to find intensities ϕ_j, Ψ_{jk} that are as general as possible but such that the equilibrium distribution of X has a tractable form.

Because our main results are still several pages away, we now jump ahead and give an example that describes the gist of our study. The next section provides a more leisurely lead-in.

Example 2.1 A Precursor to the Main Result

Suppose the network described above has a node j_0 whose service rate is independent of the other nodes; that is, for each $j \neq j_0$ and n , the $\phi_{j_0}(n)$ is independent of n_j and $\phi_j(n)$ is independent of n_{j_0} . This restriction is only used here to simplify some notation; it is not imposed in our main results. Assume that, for each n and $j, k \neq j_0$,

$$\phi_j(n) \phi_k(n - e_j) = \phi_k(n) \phi_j(n - e_k). \quad (5)$$

In addition, assume that the intensities Φ_{jk} are symmetric over a family of routing-balance partitions as described in Section 4. This routing-balance symmetry is central to our development. Then the equilibrium distribution of the network process X is $\pi(n) = c \Phi(n) \prod_j w_j^{n_j}$, where

$$\Phi(n) = \prod_{j \neq j_0} \prod_{r=1}^{n_j} \phi_j(n - \sum_{s=1}^{j-1} n_s e_s - r e_j)^{-1} \prod_{r=1}^{n_{j_0}} \phi_{j_0}(r)^{-1}. \quad (6)$$

This is a special case of our Theorem 4.6. Assumption (5) ensures that this product (6) of $\phi_1, \dots, \phi_{m-1}$ with decreasing vector arguments can be taken in any order. This is a key feature of our processes.

3 Preliminary Examples and Results

In this section, we present examples and corollaries of our results that are related to previous studies, and point out some issues that our work addresses. Our aim is to describe several types of networks that highlight a spectrum of networks with interacting nodes. We use the notation presented above. One can read this section before or after the main results in

Section 4. A reader who is familiar with network processes and prefers to see examples after the main results might want to skip ahead to Section 4.

Example 3.1 Jackson Networks

Suppose the network process X has transition rates

$$q(n, T_{jk}n) = \lambda_{jk}\phi_j(n_j),$$

where $\phi_j(n_j)$ is the service rate at node j depending on only n_j . This is the classical Jackson network with equilibrium distribution

$$\pi(n) = c\Phi(n) \prod_j w_j^{n_j},$$

where

$$\Phi(n) = \prod_j \prod_{r=1}^{n_j} \phi_j(r)^{-1}.$$

Here the non-interacting nodes give rise to a product-form equilibrium distribution. It would be informative to find a characterization of the general rates (1) for interacting nodes that also yield a product-form distribution. \square

Dependencies in a network arise naturally via subsets of nodes that have special interactive features based on some common property. For instance, subsets of work stations (nodes) in a manufacturing network may represent work centers, as described in [6], whose intra-center routing rules differ from the inter-center routing rules, and the processing rate at a station may depend on the number of units in the center in which it resides. The following is one of the simplest dependencies between subsets of nodes.

Example 3.2 Networks with Generalized Product-Form Equilibrium Distributions

Suppose there is a partition \mathcal{P} of the nodes $1, \dots, m$ based on some of their characteristics (\mathcal{P} is a collection of disjoint sets whose union is $\{1, \dots, m\}$). Assume that the transition rates of X are

$$q(n, T_{jk}n) = \lambda_{jk}\phi_J(\sum_{\ell \in J} n_\ell) \quad \text{for } j \in J, J \in \mathcal{P}.$$

That is, the departure intensity at each node $j \in J$ is a function of the total number of units in J (e.g. one processor may be serving all units in J simultaneously). It follows, by verifying the balance equations (3), that the equilibrium distribution of X is

$$\pi(n) = \prod_{J \in \mathcal{P}} \Phi_J(\sum_{j \in J} n_j) \prod_{j \in J} w_j^{n_j}$$

where $\Phi_J(\nu) = \prod_{r=1}^{\nu} \phi_J(r)^{-1}$. This distribution is a product of functions of $\{n_j : j \in J\}$ for the disjoint J in \mathcal{P} . Our results in Section 4 shed light on the question: Is this generalized-product-form valid for departure intensities ϕ_J that are more general functions of $\{n_j : j \in J\}$, or for more general transition rates? \square

Many dependencies in networks over a partition or subnetworks can be modeled by existing processes simply by ingenuity of notation. One approach is to view the subnetworks as a coarser network, as in [29], in which each subnetwork J is a “node” with a state $\{n_j : j \in J\}$. Another approach is to attach labels to the units to form a multi-type network process. Our concern will be dependencies, including those involving overlapping subsets of nodes, that cannot be handled as easily.

Another major source of network dependencies is between the routing of units and the processing rates. Our main focus will be on characterizing such dependencies in terms of how the routing intensities λ_{jk} interact with the departure-transfer intensities $\phi_j(n) \Psi_{jk}(n)$. The next two examples set the stage for this.

Example 3.3 Kelly-Whittle Networks

Suppose the network process X has transition rates

$$q(n, T_{jk}n) = \lambda_{jk} \Phi(n - e_j) / \Phi(n), \quad (7)$$

where $\Phi : S_{N-1} \cup S_N \rightarrow R^+$ and $R^+ = (0, \infty)$. Note that the Jackson network process is a special case. One can view $V(n) \equiv -\log \Phi(n)$ as the system “potential” when in state n . Then the departure intensity $\phi_j(n) = \exp[V(n) - V(n - e_j)]$ is a function of the difference in potential between the state n and n with one less unit at node j . It follows, by verifying (3), that the equilibrium distribution for this network is $\pi(n) = c \Phi(n) \prod_j w_j^{n_j}$. This process is discussed in [16] and [29].

The following result describes a variation of this process. One can prove this directly by appeal to (3); it is also a corollary of Theorem 4.6 below.

Proposition 3.4 *Suppose the transition rates of X are $q(n, T_{jk}n) = \lambda_{jk} \phi_j(n)$*

and there exists a $\Phi : S_N \rightarrow R^+$ such that

$$\Phi(n+e_j)\phi_j(n+e_j) = \Phi(n+e_k)\phi_k(n+e_k) \text{ for each } j \neq k \text{ and } n. \quad (8)$$

Then the equilibrium distribution of X is $\pi(n) = c\Phi(n)\prod_j w_j^{n_j}$. The condition (8) holds if and only if each ϕ_j is of the form

$$\phi_j(n) = \gamma(n-e_j)/\Phi(n) \quad (9)$$

for some $\gamma : S_{N-1} \rightarrow R^+$.

This result establishes that the $\Phi(n-e_j)$ in the rates (7) can be replaced by any function $\gamma(n-e_j)$ and the π remains the same. That is, π is "insensitive" to γ . (This is a generic use of the word insensitive that differs from its traditional use in queueing theory.) We initially thought this insensitivity was remarkable, but our study shows that it is a rather common phenomenon. This is because the γ "factors out" of the balance equations (3) since $\gamma(T_{jk}n - e_k) = \gamma(n-e_j)$. The throughputs at the nodes and other performance parameters of the network as shown below, however, generally depend on γ . Proposition 3.4 shows that the departure intensities in (7) and (9) are rather natural. Note that, for the process X , the π is essentially determined upon specifying Φ in (7) or (9). One of our goals is to develop the forward approach of starting with the service rates ϕ_j as the initial data and then deriving Φ and hence π as an explicit function of the ϕ_j 's. \square

Example 3.5 Kingman Reversible Networks

We say that the network process X is *reversible* if its equilibrium distribution π satisfies the *detailed balance equations*

$$\pi(n)q(n, T_{jk}n) = \pi(T_{jk}n)q(T_{jk}n, n) \text{ for each } j, k, n. \quad (10)$$

The standard definition of reversibility, as for instance in [16], requires the additional condition that X is stationary, but we do not need this. Similarly, we say that the *routing* is *reversible* if there are positive w_1, \dots, w_m that satisfy

$$w_j\lambda_{jk} = w_k\lambda_{kj} \text{ for each } j, k. \quad (11)$$

A classic example of such routing is a star-like network with node 1 as the center, and the $\lambda_{1j}, \lambda_{j1}$ are positive for $j = 2, \dots, m$ and $\lambda_{jk} = 0$ otherwise. The following result is discussed in [17] and [29]; it follows immediately

from (10) and is a corollary of Theorem 4.6. Here we use $\Lambda_{jk}(n)$ to denote the departure-transfer rate $\phi_j(n)\Psi_{jk}(n)$.

Proposition 3.6 *Suppose the network process X has transition rates*

$$q(n, T_{jk}n) = \lambda_{jk}\Lambda_{jk}(n) \quad (12)$$

and the routing is reversible. Then X is reversible if and only if each Λ_{jk} is of the form

$$\Lambda_{jk}(n) = \Phi(n - e_j)\Phi(n)^{-1}\gamma_{jk}(n - e_j) \quad (13)$$

for some $\Phi : S_{N-1} \cup S_N \rightarrow R^+$ and $\gamma_{jk} : S_{N-1} \rightarrow R^+$ that satisfy $\gamma_{jk} = \gamma_{kj}$. In this case, the equilibrium distribution of X is $\pi(n) = c\Phi(n) \prod_j w_j^{n_j}$.

In expression (13), the $\Phi(n - e_j)$ can be incorporated in $\gamma_{jk}(n - e_j)$, but the given form is more convenient. Although the function Φ is typically defined on only S_N , its definition usually extends readily to the larger domain $S_1 \cup \dots \cup S_N$. An example of this result is a network with rates

$$q(n, T_{jk}n) = \lambda_{jk}\phi_j(n_j)\psi_k(n_k),$$

and

$$\Phi(n) = \prod_j \prod_{r=1}^{n_j} \psi_j(r-1)/\phi_j(r)$$

$$\gamma_{jk}(n) = \psi_j(n_j)\psi_k(n_k).$$

The $\psi_k(n_k)$ can be viewed as the intensity at which node k attracts or accepts units from any other node. Other types of general intensities (13) are $1/\Phi(n)$, $\gamma(n - e_j, n + e_k)/\Phi(n)$ and $\gamma(n - e_j, n + e_k)$ (here $\Phi \equiv 1$). See [5], [6], [18], and [24] for related examples. \square

The two preceding examples represent extreme cases of system-dependent transition rates. In Example 3.3 the networks have general routing rates λ_{jk} but the departure-transfer rates are restricted. On the other hand, the networks in Example 3.5 have general departure-transfer rates but the routing rates are restricted to be reversible. We shall study a spectrum of networks between these cases that have varying degrees of routing reversibility. The following is one such example.

Example 3.7 Networks with Reversible Routing Between (or Within) Disjoint Subnetworks

Suppose the network is comprised of a union of disjoint subnetworks (i.e.

subsets of nodes), and let S_j denote the subnetwork containing node j . Assume that the routing of units between any two subnetworks is reversible in that the w_1, \dots, w_m that satisfy (2) also satisfy

$$w_j \lambda_{jk} = w_k \lambda_{kj} \text{ for each } k \in S_j^c \quad (14)$$

where $S_j^c \equiv \{1, \dots, m\} \setminus S_j$. Now, assume that the transition rates of X are

$$q(n, T_{jk}n) = \begin{cases} \lambda_{jk} \Phi(n - e_j) \Phi(n)^{-1} \gamma_{jk}(n - e_j) & k \in S_j^c \\ \lambda_{jk} \Phi(n - e_j) \Phi(n)^{-1} \gamma_{S_j}(n - e_j) & k \in S_j. \end{cases} \quad (15)$$

Here $\gamma_S : S_{N-1} \rightarrow R^+$, and $\gamma_{jk} : S_{N-1} \rightarrow R^+$ are such that $\gamma_{jk} = \gamma_{kj}$. In this network, the movement of units between subnetworks is comparable to the reversible movement in Example 3.5 and the movement of units within each subnetwork is comparable to the movement of units in Example 3.3. It follows from Theorem 4.6 that the equilibrium distribution of X is $\pi(n) = c \Phi(n) \Pi_j w_j^{n_j}$. There is an analogous network process in which the routing of units within each subnetwork is reversible and the routing between subnetworks need not be. This is defined as above by simply reversing the roles of S_j^c and S_j in expressions (14), (15). These examples exhibit a partial routing reversibility in terms of only disjoint subnetworks. We will also discuss related partial reversibilities via non-disjoint subnetworks. \square

The examples presented above suggest that they belong to a wider family of processes that apparently have a convenient canonical form. This is the subject we shall now develop.

4 A Canonical Network Process

Throughout this section we assume that $X = \{X_t : t \geq 0\}$ is a closed Markovian network process with transition rates

$$q(n, T_{jk}n) = \lambda_{jk} \phi_j(n) \Psi_{jk}(n).$$

We shall characterize the equilibrium distribution of X for rather general intensity functions ϕ_j, Ψ_{jk} .

We begin with a few definitions. We have seen that partial reversibility or balance of the routing rates λ_{jk} on subsets of nodes allows more general

intensities ϕ_j , Ψ_{jk} for the nodes. A convenient way of characterizing partial balance properties of the routing rates is by routing-balance partitions defined as follows.

Definition 4.1 Let w_1, \dots, w_m be positive real numbers that satisfy the routing balance equations

$$\sum_k (w_j \lambda_{jk} - w_k \lambda_{kj}) = 0 \quad j = 1, \dots, m. \quad (16)$$

For each j , let D_j denote the set of all nodes k that satisfy the “detailed balance” condition

$$w_j \lambda_{jk} = w_k \lambda_{kj}. \quad (17)$$

Let \mathcal{B}_j be any collection of disjoint sets of nodes in $D_j^c \equiv \{1, \dots, m\} \setminus D_j$ that partition D_j^c and satisfy the “partial balance” condition

$$\sum_{k \in B} (w_j \lambda_{jk} - w_k \lambda_{kj}) = 0 \quad \text{for each } B \in \mathcal{B}_j. \quad (18)$$

We call $\mathcal{B}_1, \dots, \mathcal{B}_m$ routing-balance partitions.

Note that the *detailed-balance set* D_j and the *partial-balance partition* \mathcal{B}_j do not depend on the particular choice of w_1, \dots, w_m . This follows since any two vectors that satisfy (16) differ by only a constant multiple, and so (17) and (18) are satisfied by any such vector. One can therefore view D_j and \mathcal{B}_j as being defined solely by $\{\lambda_{jk}\}$. The D_j and \mathcal{B}_j exist for any $\{\lambda_{jk}\}$. Clearly D_j is a unique set. Two extreme cases are $D_j = \{j\}$, which represents no detailed balance, and $D_j = \{1, \dots, m\}$, which represents complete detailed balance implying that \mathcal{B}_j is the empty family. Note that $D_1 = \dots = D_m = \{1, \dots, m\}$ if and only if the routing is reversible. Clearly D_j contains all k such that $\lambda_{jk} = \lambda_{kj} = 0$.

Keep in mind that \mathcal{B}_j is any partition of D_j^c and hence is not unique. The coarsest partition $\mathcal{B}_j \equiv \{D_j^c\}$ is always a possibility. Note that each (nonempty) set B in \mathcal{B}_j consists of at least two nodes; if B were to consist of only one node, then that node would be in D_j . As we shall soon see, finer partitions \mathcal{B}_j allow more general service-transfer intensities. Therefore, when there are several possibilities for \mathcal{B}_j (which is not uncommon), then one should choose a fine one. It would be nice if there were a unique “finest” partition \mathcal{B}_j , but there generally isn’t; see the Appendix. We already used routing-balance partitions in Example 3.7; there $D_j = S_j^c$ and $\mathcal{B}_j = \{S_j\}$.

Note that each network process has routing-balance partitions. Our definition does not represent a new property –it only labels the partial balance that is implicit in the network routing. The routing-balance partitions may be a prominent feature of the network due to its design, such as Example 3.7, but in other cases they may be inconspicuous.

Routing-balance partitions are important for demarcating the partial balance of the entire network process as follows.

Definition 4.2 Let $\mathcal{B}_1, \dots, \mathcal{B}_m$ be routing-balance partitions for the process X . Suppose that π is a probability measure on S with positive values such that, for each j and n ,

$$\pi(n)q(n, T_{jk}n) = \pi(T_{jk}n)q(T_{jk}n, n) \quad k \in D_j \quad (19)$$

$$\sum_{k \in B} [\pi(n)q(n, T_{jk}n) - \pi(T_{jk}n)q(T_{jk}n, n)] = 0 \quad B \in \mathcal{B}_j. \quad (20)$$

These partial balance conditions imply the total balance conditions (9), and hence π is the equilibrium distribution for X . We say that X is balanced over $\mathcal{B}_1, \dots, \mathcal{B}_m$.

To see the meaning of this definition, first recall that $\pi(n)q(n, T_{jk}n)$ is the number of occurrences per unit time of the event that a unit moves from j to k when X is in state n . The Appendix contains a concise description of such occurrence rates. The $\pi(n)q(n, T_{jk}n)$ is also called the rate of flow (or probability flux) from j to k when X is in state n . Then

$$\sum_{n \in S} \pi(n)q(n, T_{jk}n)$$

is the rate of flow of units from j to k . Note that condition (19) implies that for each j and k between which the routing is reversible as in (17), the rate of flow from j to k equals the rate of the reverse flow from k to j .

Similarly, condition (20) implies that the rate at which units flow from j into B equals the rate of flow of units from B into j . The network studies to date have emphasized the partial balance condition (20) only for the case in which $B = \{1, \dots, m\}$. Conditions (19), (20) simply bring to light further micro-level insights into the balance of the flows.

We now introduce a key concept related to partial balance.

Definition 4.3 Functions $\gamma_{jk} : S_N \rightarrow [0, \infty)$, $j, k = 1, \dots, m$, are symmetric over the routing balance partitions $\mathcal{B}_1, \dots, \mathcal{B}_m$ if

$$\gamma_{jk}(n + e_j) = \gamma_{kj}(n + e_k) \quad \text{for each } j, k, n \quad (21)$$

and

$$\gamma_{jk} = \gamma_{j\ell} \text{ for each } k, \ell \in B, B \in \mathcal{B}_j \text{ and } j. \quad (22)$$

This simply says that the functions $\gamma_{jk}^*(n) \equiv \gamma_{jk}(n+e_j)$ are such that $\gamma_{jk}^* = \gamma_{kj}^*$, and γ_{jk}^* is the same for each $k \in B$, where $B \in \mathcal{B}_j$. An example is

$$\gamma_{jk}(n) = \begin{cases} g_{jk}(n - e_j, n + e_k) & k \in D_j \\ h_{jB}(n - e_j, n + e_k) & k \in B, B \in \mathcal{B}_j, \end{cases}$$

where g_{jk} and h_{jB} are functions on $S_{N-1} \times S_{N+1}$ and $g_{jk} = g_{kj}$. Note that this notion of routing-symmetric functions is not related to the notion of a symmetric queue.

Routing-symmetric functions play an important role in our analysis. They are useful for linking the service-transfer intensities to the routing intensities, which we will do shortly. A more fundamental use of routing-symmetric functions relates to partial balance as follows. Here we use $\Lambda_{jk}(n)$ to represent the departure-transfer intensity $\phi_j(n)\Psi_{jk}(n)$.

Lemma 4.4 *Let X denote the network process with transition rates $q(n, T_{jk}n) = \lambda_{jk}\Lambda_{jk}(n)$. Suppose there is a function $\Phi : S \rightarrow R^+$ such that the functions $\Phi(n)\Lambda_{jk}(n)$ are symmetric over the routing-balance partitions $\mathcal{B}_1, \dots, \mathcal{B}_m$. Then the equilibrium distribution of X is $\pi(n) = c\Phi(n) \prod_j w_j^{n_j}$, and X is balanced over $\mathcal{B}_1, \dots, \mathcal{B}_m$.*

Proof. It suffices to show that the specified π satisfies the partial balance equations (19), (20). Fix n and j . If $n_j = 0$, then (19), (20) are trivially satisfied. Hereafter, assume $n_j > 0$. From the definition of π and since $\Phi\Lambda_{jk}$ are symmetric over $\mathcal{B}_1, \dots, \mathcal{B}_m$, it follows that, for each k ,

$$\pi(T_{jk}n) = \pi(n)\Phi(T_{jk}n)\Phi(n)^{-1}w_k w_j^{-1}, \text{ and,}$$

$$\Phi(T_{jk}n)\Lambda_{kj}(T_{jk}n) = \Phi(n)\Lambda_{jk}(n).$$

Applying these identities, we have

$$\begin{aligned} \pi(T_{jk}n)q(n, T_{jk}n) &= \pi(n)\Phi(T_{jk}n)\Phi(n)^{-1}w_k w_j^{-1}\lambda_{kj}\Lambda_{kj}(T_{jk}n) \\ &= \pi(n)w_k w_j^{-1}\lambda_{kj}\Lambda_{jk}(n). \end{aligned} \quad (23)$$

For $k \in D_j$, we know that $w_k w_j^{-1}\lambda_{kj} = \lambda_{jk}$, and so (23) immediately yields (19). Also, for $B \in \mathcal{B}_j$ and a fixed k' in B , it follows from (23) and $\Lambda_{jk} = \Lambda_{jk'}$,

$k \in B$, that

$$\begin{aligned} \sum_{k \in B} \pi(T_{jk}n) q(T_{jk}n, n) &= \pi(n) \Lambda_{jk}(n) w_j^{-1} \sum_{k \in B} w_k \lambda_{kj} \\ &= \pi(n) \Lambda_{jk}(n) \sum_{k \in B} \lambda_{jk} \\ &= \pi(n) \sum_{k \in B} q(n, T_{jk}n). \end{aligned}$$

This proves (20). \square

The following is a consequence of Lemma 4.4.

Criterion for Determining an Equilibrium Distribution 4.5 *Suppose π is a distribution on S with positive values such that the functions*

$$\gamma_{jk}(n) = \pi(n) w_j^{-1} \Lambda_{jk}(n)$$

are symmetric over B_1, \dots, B_m . Then π is the equilibrium distribution of X , and X is balanced over B_1, \dots, B_m .

Proof. Let $W(n) = \prod_j w_j^{n_j}$ and define $\Phi(n) = \pi(n)/W(n)$. It suffices, by Lemma 4.4, to show that the functions

$$\Phi(n) \Lambda_{jk}(n) = \gamma_{jk}(n) w_j W(n)^{-1}.$$

are symmetric over B_1, \dots, B_m . But this follows since the γ_{jk} are symmetric by assumption and so is $w_j W(n)^{-1}$ since $w_j W(n + e_j)^{-1} = W(n)$, independent of j . \square

This criterion is surprisingly simple but very useful. It says that to verify that a distribution π satisfies (19), (20) and hence is the equilibrium distribution of X , one need only verify that the specified function γ_{jk} satisfy the simple symmetry conditions (21), (22).

We are now ready to present our main result. We will use the following condition: There is a $j_0 \in \{1, \dots, m\}$ such that, for each n and $j, k \neq j_0$,

$$\phi_j(n) \phi_k(n - e_j + e_{j_0}) \phi_{j_0}(n - e_k + e_{j_0}) = \phi_k(n) \phi_j(n - e_k + e_{j_0}) \phi_{j_0}(n - e_j + e_{j_0}). \quad (24)$$

In addition, we define, for $n \in S_N$,

$$\Phi(n) = \prod_{\nu=1}^M \phi_{j_0}(n - \sum_{s=1}^{\nu} e_{j_s} + \nu e_{j_0}) / \phi_{j_\nu}(n - \sum_{s=1}^{\nu-1} e_{j_s} + (\nu-1)e_{j_0}), \quad (25)$$

where $M = N - n_{j_0}$ and j_1, \dots, j_M is a sequence in $\{1, \dots, m\} \setminus \{j_0\}$ such that $\sum_{s=1}^{\nu} e_{j_s} \leq n$ (componentwise) for each ν . Condition (24) ensures that (25) is the same for any such sequence j_1, \dots, j_M ; we justify this in Lemma 4.7.

Theorem 4.6 *Suppose the network process X has transition rates*

$$q(n, T_{jk}n) = \lambda_{jk} \phi_j(n) \Psi_{jk}(n)$$

where ϕ_1, \dots, ϕ_m satisfy condition (24) and the Ψ_{jk} are symmetric over the routing-balance partitions $\mathcal{B}_1, \dots, \mathcal{B}_m$. Then the equilibrium distribution of X is $\pi(n) = c \Phi(n) \prod_j w_j^{n_j}$, where Φ is given by (25), and X is balanced over $\mathcal{B}_1, \dots, \mathcal{B}_m$.

Proof. It suffices, by Lemma 4.4, to show that the functions

$$\gamma_{jk}(n) \equiv \Phi(n) \phi_j(n) \Psi_{jk}(n)$$

are symmetric over $\mathcal{B}_1, \dots, \mathcal{B}_m$. From Lemma 4.7(c) below, we know that condition (24) ensures that $\phi_j(n+e_j) \Phi(n+e_j)$ is independent of j . This and the symmetry of Ψ_{jk} over $\mathcal{B}_1, \dots, \mathcal{B}_m$ imply that γ_{jk} are symmetric over $\mathcal{B}_1, \dots, \mathcal{B}_m$. \square

A major step in the preceding proof is justified by the next result, which also explains the meaning of the condition (24). In fact, the equivalence of statements (a) and (d) here is the heart of Theorem 4.6.

Lemma 4.7 *The following statements are equivalent.*

- (a) *The functions ϕ_1, \dots, ϕ_m satisfy condition (24).*
- (b) *The product (25) is the same for any sequence j_1, \dots, j_M with values in $\{1, \dots, m\} \setminus \{j_0\}$ such that $\sum_{s=1}^{\nu} e_{j_s} \leq n$ for each ν .*
- (c) *There exists a function $\Phi' : S_N \rightarrow R^+$ such that for each $n \in S_{N-1}$, the $\Phi'(n+e_j) \phi_j(n+e_j)$ is independent of j .*
- (d) *There is a function $\Phi^* : S_{N-1} \cup S_N \rightarrow R^+$ such that*

$$\phi_j(n) = \Phi^*(n - e_j + e_{j_0}) \Phi^*(n)^{-1} \phi_{j_0}(n - e_j + e_{j_0}) \quad \text{for each } n \text{ and } j. \quad (26)$$

If these statements are satisfied, then Φ' and Φ^ are constant multiples of Φ .*

Proof. Suppose (a) holds. We will prove (b) by induction on $M = 0, \dots, N$. Let $\Phi_M(n)$ denote the product (25). Clearly (b) is trivially satisfied for $M = 0$ since $\Phi_0(n) = 1$. Now assume (b) is true up to some $M < N$. For $n \in S_N$ with $N - n_{j_0} = M + 1$, let $\Phi_{M+1}(n)$ and $\Phi_{M+1}^*(n)$ denote the product (25) for two sequences j_1, \dots, j_{M+1} and j_1^*, \dots, j_{M+1}^* . Our aim is to show that $\Phi_{M+1}(n) = \Phi_{M+1}^*(n)$. Factoring out the first term in each of these products, we can write them as

$$\Phi_{M+1}(n) = \phi_{j_0}(n - e_{j_1} + e_{j_0})\phi_{j_1}(n)^{-1}\Phi_M(n - e_{j_1} + e_{j_0}) \quad (27)$$

$$\Phi_{M+1}^*(n) = \phi_{j_0}(n - e_{j_1^*} + e_{j_0})\phi_{j_1^*}(n)^{-1}\Phi_M(n - e_{j_1^*} + e_{j_0}).$$

For the last term we used the induction hypothesis that $\Phi_M^* = \Phi_M$. If $j_1 = j_1^*$, then the two preceding expressions are the same. Now suppose $j_1 \neq j_1^*$. Under the induction hypothesis, $\Phi_M(n - e_{j_1} + e_{j_0})$ can be expressed as (25) for a sequence beginning with j_1^* , since $j_1 \neq j_1^*$ ensures that $n_{j_1^*} \geq 1$. Factoring out the first term in this Φ_M , expression (27) becomes

$$\begin{aligned} \Phi_{M+1}(n) &= \phi_{j_0}(n - e_{j_1} + e_{j_0})\phi_{j_1}(n)^{-1}\phi_{j_0}(n - e_{j_1} - e_{j_1^*} + 2e_{j_0}) \\ &\quad \phi_{j_1^*}(n - e_{j_1} + e_{j_0})\Phi_{M-1}(n - e_{j_1} - e_{j_1^*} + 2e_{j_0}). \end{aligned}$$

By similar reasoning,

$$\begin{aligned} \Phi_{M+1}^*(n) &= \phi_{j_0}(n - e_{j_1} + e_{j_0})\phi_{j_1^*}(n)^{-1}\phi_{j_0}(n - e_{j_1} - e_{j_1^*} + 2e_{j_0}) \\ &\quad \phi_{j_1}(n - e_{j_1^*} + e_{j_1})\Phi_{M-1}(n - e_{j_1} + e_{j_1^*} + 2e_{j_0}). \end{aligned}$$

Under the assumption (24), these two expressions are equal, and so the induction is complete. Thus we have established that (a) implies (b).

If (b) holds, then expression (25) ensures that, for each $j \neq j_0$ and $n \in S_{N-1}$,

$$\Phi(n + e_j)\phi_j(n + e_j) = \Phi(n + e_{j_0})\phi_{j_0}(n + e_{j_0}).$$

Thus (c) is satisfied with $\Phi' = \Phi$. Next, observe that (c) obviously implies (d) with $\Phi^* = \Phi'$. Finally, (d) implies (a), since the ϕ_j specified in (d) clearly satisfy (24). This finishes the proof that (a) - (d) are equivalent.

Now, suppose these statements are true. Then from (d) and iterations on j_1, \dots, j_m , we have

$$\Phi^*(n) = \phi_{j_0}(n - e_{j_1} + e_{j_0})\phi_{j_1}(n)^{-1}\Phi^*(n - e_{j_1} + e_{j_0})$$

$$\begin{aligned}
&= \phi_{j_0}(n - e_{j_1} + e_{j_0})\phi_{j_1}(n)^{-1}\phi_{j_0}(n - e_{j_1} - e_{j_2} + 2e_{j_0}) \\
&\quad \phi_{j_2}(n - e_{j_1} + e_{j_0})^{-1}\Phi^*(n - e_{j_1} - e_{j_2} + 2e_{j_0}) \\
&\quad \dots \\
&= \Phi^*(Ne_{j_0})\Phi(n).
\end{aligned}$$

Thus Φ^* is a multiple of Φ . Also, we noted above that (c) implies (d) with $\Phi' = \Phi^*$. Thus Φ' is a multiple of Φ . \square

Theorem 4.6 describes a canonical network process with system-dependent transition rates. Its assumptions are easy to check for an actual network with specified ϕ_j and Ψ_{jk} . Each of the examples and results in the preceding sections is a special case of Theorem 4.6. In Section 5, we give further examples and discuss canonical networks with blocking (Theorem 5.1). Many networks that have been of interest to computer scientists and operations research analysts will probably fit this canonical form, or its modification with blocking or multiple types of units that we discuss shortly. A recent example is [25], which uses an indirect assumption resembling (b) in Lemma 4.7. For those networks that do not fit, one may still be able to use Criterion 4.5 to determine their equilibrium distributions. Other approaches that have been used to analyze dependencies in networks involve job local balance and adjoint processes [12], [13], and coupling of nodes to form networks [22] and [29]. These are broad approaches or recipes, like Lemma 4.4 and Criterion 4.5, that might be useful for deriving equilibrium distributions not covered by Theorems 4.6 and 5.1.

Remarks 4.8

- (a) The equilibrium distribution in Theorem 4.6 is a tractable function of the ϕ'_j 's that is a natural generalization of the Jackson equilibrium distribution. The Φ is an interchangeable product of the ϕ'_j 's and the vector $n - \sum_{s=1}^{\nu} e_{j_s} + \nu e_{j_0}$ is increasing in the j_0 th coordinate and is decreasing in the other coordinates as ν increases. The product (25) is the same for any j_1, \dots, j_M , and so, for the case in which this sequence is increasing, (25) becomes

$$\Phi(n) = \prod_{j \neq j_0} \prod_{r=1}^{n_j} \phi_{j_0}(h_j(n) - re_j + re_{j_0}) / \phi_j(h_j(n) - (r-1)e_j + (r-1)e_{j_0})$$

where

$$h_j(n) = (0, \dots, 0, n_j, n_{j+1}, \dots, n_m) + \left(\sum_{s=1}^{j-1} n_s \right) e_{j_0}.$$

- (b) Condition (24) and the product (25) reduce to (5) and (6) in Example 2.1 if j_0 is a node such that for each $j \neq j_0$ and n , the $\phi_{j_0}(n)$ is independent of n_j and $\phi_j(n)$ is independent of n_{j_0} . A similar simplification occurs for open networks when j_0 is the outside node; see Theorem 7.2.
- (c) Recall that in the Kelly-Whittle networks in Example 3.3, the departure rates ϕ_j are “assumed” to be of the form (26). From the equivalence of (a) and (d) in Lemma 4.7, we now know an easily checkable condition on the basic data ϕ_j , namely (24), that leads to the form (26).
- (d) The equilibrium distribution π does not depend on Ψ_{jk} . This insensitivity follows since, by assumption, $\Psi_{jk}(n+e_j) = \Psi_{kj}(n+e_k)$ and hence this term factors out of the rates in (19), (20). The throughputs and other system parameters, however, generally depend on Φ_{jk} .
- (e) If $\lambda_{jk} = 0$ and hence $\lambda_{jk}\phi_j(n)\Psi_{jk}(n) = 0$, then ϕ_j and Ψ_{jk} for that particular j, k are irrelevant and can therefore be defined arbitrarily. Consequently, any assumption on them is not a restriction.
- (f) Whether or not π has a generalized product form as in Example 3.2 is a secondary issue. Such a form arises when there is a partition \mathcal{P} of $\{1, \dots, m\}$ such that $\phi_j(n)$ is a function of only $\{n_k : k \in S_j\}$, where S_j is the set in \mathcal{P} that contains j . The partition will usually be apparent from the particular way in which the nodes interact as “families” of nodes.

Remark 4.9 Normalizing Constants

Although the normalizing constant c defined by

$$c^{-1} = \sum_{n \in S_N} \Phi(n) \prod_{j=1}^n w_j^{n_j}$$

can be computed by total enumeration for small networks, a more parsimonious approach would generally be required for large networks. Computing

c is as hard as computing a normalization constant for a Gibb's distribution for a Markov random field. The difficulty, of course, would depend on the complexity of $\Phi(n)$. There is apparently no "universal" procedure that would be efficient for all Φ 's. The worst case would be that the values of $\Phi(n)$ are nearly all different and have no pattern (like a table of arbitrary numbers). These worst cases would be the exception rather than the rule since actual networks have natural dependencies that lead to a highly structured $\Phi(n)$. For particular applications, ad hoc computational procedures or approximations based on the structure of $\Phi(n)$ would be in order.

Remark 4.10 Throughputs

For the process in Theorem 4.6, the throughput from node j to node k is

$$\begin{aligned} r_{jk} &= \lim_{t \rightarrow \infty} t^{-1} \sum_{s \leq t} 1(X_s = X_{s-} - e_j + e_k) \\ &= \sum_{n \in S_N} \pi(n) q(n, T_{jk}n) \\ &= c \lambda_{jk} \sum_{n \in S_N} [\Phi(n) \phi_j(n) \Psi_{jk}(n) \prod_{\ell} w_{\ell}^{n_{\ell}}]. \end{aligned}$$

The Appendix gives further insights on throughputs. Factoring out a w_j from the last product, we can write this as

$$r_{jk} = w_j \lambda_{jk} (c/c_{N-1}) \sum_{n \in S_{N-1}} \pi_{N-1}(n) \gamma(n) \Psi_{jk}(n + e_j) / \Phi(n)$$

where $\gamma(n) \equiv \Phi(n + e_{j_0}) \phi_{j_0}(n + e_{j_0})$ and

$$\pi_{N-1}(n) = c_{N-1} \Phi(n) \prod w_j^{n_j} \quad n \in S_{N-1}$$

is the equilibrium distribution of the network process X with $N-1$ units in it. This expression for r_{jk} will usually simplify depending on the structure of Ψ_{jk} and Φ , and this form might be convenient for recursive computations for $N = 1, 2, \dots$. As with normalizing constants, particular applications for large networks would warrant ad hoc computational procedures.

Remark 4.11 Multiple Types of Units

The results herein readily extend to networks with multiple types of units. A standard way of doing this is to append another letter to the node numbers as follows. Let $n_{\alpha j}$ denote the number of type α units at node j , where $\alpha \in C$, a set of types or classes. A generic state of the network processes X

would therefore be $n = \{n_{\alpha j} : \alpha \in C, j = 1, \dots, m\}$. A typical transition would be from n to $T_{\alpha j, \beta k} n$ which means that a type α unit at node j enters node k as a type β unit. The rate of such a transition would be

$$q(n, T_{\alpha j, \beta k} n) = \lambda_{\alpha j, \beta k} \phi_{\alpha j}(n) \Psi_{\alpha j, \beta k}(n).$$

In an obvious manner, one can restate each result herein in terms of the doubly subscripted variables $\alpha j, \beta k, \dots$ instead of simply j, k, \dots

5 Blocking of Transitions

One of the commonest system dependencies in a network is that of blocking of transitions due to certain constraints. The simplest example is that the number of units at a node (or a set of nodes) cannot exceed a prescribed value and if a unit attempts to enter the node when the number of units there equals that value, then it is blocked - the unit remains where it is for another service cycle. This is often called communication blocking. Other examples of constraints that result in blocked transitions are as follows.

- The service rate at a node is zero if the number of units there is below a prescribed value.
- The entry of a unit from node j into a set of nodes K is suppressed if the number of units in K exceeds the number in j .
- Each node must contain at least one unit.

These go/no-go constraints are simply restrictions on the usual transitions of the network process that confine it to a smaller state space. We purposely avoided these constraints in the preceding, by assuming that $\phi_j(n) \Psi_{jk}(n)$ is positive, so as not to cloud the main ideas. We now show how such constraints can be incorporated into our canonical network.

Suppose that $X = \{X_t : t \geq 0\}$ is the closed network process described in Theorem 4.6. We shall consider a modification of this process in which the state space is restricted to a subset S^* of S_N . Specifically, we let $X^* = \{X_t^* : t \geq 0\}$ be a Markov network process with transition rates

$$q^*(n, T_{jk} n) = \lambda_{jk} \phi_j(n) \Psi_{jk}(n) \alpha_{jk}(n) \quad n \in S^*. \quad (28)$$

Here $\alpha_{jk}(n) \geq 0$ is an additional intensity that suppresses certain transitions when it is zero. The process X^* evolves just as X would subject to the modification that if X were to have a transition from n to $T_{jk}n$ in S^* and $\alpha_{jk}(n) = 0$, then this transition is suppressed. The interpretation is that the unit due to move from j to k returns to j to endure another sojourn there, as a new entry would. In some applications, the $\alpha_{jk}(n)$ might be the probability that the potential movement of a unit from j to k is accepted. This formulation of blocking via (28) covers a variety of system-dependent blockings that have not been characterized before.

Since X^* is similar to X on the smaller state space S^* , it is natural to ask whether the equilibrium distribution π^* for X^* is a truncation of the distribution π for X . The next result addresses the issue. For convenience, we assume that X^* is irreducible on the space S^* .

Theorem 5.1 *If α_{jk} are symmetric over the routing balance partitions B_1, \dots, B_m , then the equilibrium distribution of X^* is*

$$\pi^*(n) = \pi(n) / \sum_{n' \in S^*} \pi(n'), \quad n \in S^*,$$

and X^* is balanced over B_1, \dots, B_m .

Proof. Let $W(n) = \prod_j w_j^{n_j}$. By Criterion 4.5, it suffices to establish that the functions

$$\begin{aligned} \gamma_{jk}(n) &= \pi^*(n) w_j^{-1} \Delta_{jk}(n) \\ &= \Phi(n) \phi_j(n) \Psi_{jk}(n) \alpha_{jk}(n) w_j^{-1} W(n) / \sum_{n' \in S^*} \pi(n') \end{aligned}$$

are symmetric over B_1, \dots, B_m . But this follows since the proof of Theorem 4.6 showed that $\Phi(n) \phi_j(n) \Psi_{jk}(n)$ satisfy this condition, the α_{jk} satisfy it by assumption, and $w_j^{-1} W(n)$ satisfy it since $w_j^{-1} W(n + e_j) = W(n)$. \square

Example 5.2 State-Dependent Acceptance Sets

Suppose X^* represents the network process X in which a transition from n to $T_{jk}n$ is accepted if and only if k is in a prescribed set $A_j(n)$ of nodes that accept units from j when in state n . Then X^* can be modeled as above with

$$\alpha_{jk}(n) = 1(k \in A_j(n)).$$

Here $1(\cdot)$ is the indicator function. These α_{jk} are symmetric over B_1, \dots, B_m if and only if the $A_j(n)$ are such that, for each j and $n \in S^*$,

$$k \in A_j(n) \text{ is equivalent to } j \in A_k(T_{jk}n), \text{ and}$$

$B \cap A_j(n) = B$ or is empty, for each $B \in \mathcal{B}_j$.

Under these conditions, the conclusions of Theorem 5.1 apply to this process X^* . \square

Example 5.3 State-Independent Acceptance Sets

A standard form of blocking is simply to consider X restricted to an arbitrary set S^* . This is modeled by X^* with

$$\alpha_{jk}(n) = 1(T_{jk}n \in S^*).$$

One can view this as Example 5.2 with $A_j(n) = \{k : T_{jk}n \in S^*\}$. In this case, the α_{jk} are symmetric over $\mathcal{B}_1, \dots, \mathcal{B}_m$ if and only if the S^* is such that, for each j and $n \in S^*$,

$$T_{jk}n \in S^* \text{ is equivalent to } \{T_{jk}n : k \in B\} \subset S^*, \quad B \in \mathcal{B}^j.$$

Therefore, if S^* satisfies this condition, then the conclusions of Theorem 5.1 apply to X^* . This type of result is known for Jackson processes. \square

Example 5.4 Maximum Capacities at Nodes with Reversible Routing

Suppose the network process X has a subset of nodes D such that

$$w_j \lambda_{jk} = w_k \lambda_{jk} \text{ for each } j, k \in D.$$

That is, the routing between each pair of nodes in D is reversible in that it satisfies the detailed balance condition. Let $D_0 = \{k \in D : \lambda_{jk} = 0, j \in D^c\}$, the set of nodes in the interior of D that cannot be reached from D^c in one transition. Suppose that each node k in D_0 can accommodate at most M_k units. Then the resulting process X^* is as in Theorem 5.1 with

$$\alpha_{jk}(n) = \begin{cases} 1(n_k + 1 \leq M_k) & j, k \in D_0 \\ 1 & \text{otherwise.} \end{cases}$$

To verify that these α_{jk} are symmetric over $\mathcal{B}_1, \dots, \mathcal{B}_m$, it is convenient to write

$$\alpha_{jk}(n) = 1(n \leq \mathbf{M}, T_{jk}n \leq \mathbf{M})$$

where $\mathbf{M} = (M_1, \dots, M_m)$ and $M_j = N$ for $j \in D_0$. The condition $n \leq \mathbf{M}$ ensures that $n \in S^*$. Then clearly $\alpha_{jk}(n + e_j) = \alpha_{kj}(n + e_k)$ for each $j \neq k$, and

$$\alpha_{jk}(n) = \alpha_{jk'}(n) = 1 \text{ for each } j \text{ and } k, k' \in D_j^c,$$

since $D \subset D_j$, when $j \in D$. The latter condition implies that α_{jk} are symmetric over “any” $\mathcal{B}_1, \dots, \mathcal{B}_m$. \square

Example 5.5 Maximum Capacities on Sets of Nodes

Suppose the network process X has a partition of nodes \mathcal{P} as in Example 3.7 such that

$$w_j \lambda_{jk} = w_k \lambda_{kj} \text{ for each } j \text{ and } k \in S_j^c,$$

where S_j is the set in \mathcal{P} that contains node j . The vector $N(n) = (\sum_{j \in S} n_j : S \in \mathcal{P})$ records the numbers of units in the sets in \mathcal{P} . Assume that each set S in \mathcal{P} can accommodate at most M_S units. In other words,

$$N(n) \leq \mathbf{M} \equiv (M_S : S \in \mathcal{P}).$$

This type of constraint is used in [6], [7], and [8]. Then the resulting network process X^* can be expressed as in Theorem 5.1 with

$$\alpha_{jk}(n) = 1(N(n) \leq \mathbf{M}, N(T_{jk}n) \leq \mathbf{M}).$$

An easy check shows that the α_{jk} are symmetric over any $\mathcal{B}_1, \dots, \mathcal{B}_m$, since

$$N(T_{jk}n) = N(T_{kj}n) = N(n) \text{ for } k, k' \in D_j^c = S_j.$$

Note that the single-node constraints as in the previous example could also be incorporated in this type of network. \square

Example 5.6 Transfer Rates That Help Equalize Congestion

Actual networks typically have flexible servicing and routing capabilities built in so as to help equalize congestion, or to maximize throughput or some other parameter. There are many conceivable transfer rate functions, such as route-to-the-shortest-queue, that, unfortunately, do not lead to tractable equilibrium distributions. The following is a simple illustration of the transfer rate features that one can model in our framework.

Consider a network in which there is a set of nodes D as in Example 5.4 such that the routing is reversible between any pair of nodes in D . The congestion in D is to be equalized to a reasonable degree. Accordingly, the network is designed so that the transition rates of X are

$$q(n, T_{jk}n) = \begin{cases} \lambda_{jk} \mu_j(n_j) \eta_j(n_j) \alpha_k(n_k) & j, k \text{ in } D \\ \lambda_{jk} \mu_j(n_j) & \text{otherwise.} \end{cases}$$

Here $\mu_j(n_j)$ is the actual service rate at node j , the $\alpha_k(n_k)$ is the intensity (or probability) at which node k accepts units and $\eta_j(n_j)$ is an expediting intensity. To level the congestion in D , the $\alpha_k(n_k)$ is decreasing in n_k and $\eta_j(n_j)$ is increasing in n_j . Furthermore, each node k in D_0 (the nodes that cannot be reached in one transition from outside D) is restricted to accommodate at most M_k units, and so $\alpha_k(n_k) = 0$ for $n_k \geq M_k$. We can write the transition rates above as

$$q(n, T_{jk}n) = \lambda_{jk} \phi_j \Psi_{jk}(n),$$

where

$$\phi_j(n) = \begin{cases} \mu_j(n_j) \eta_j(n_j) / \alpha_j(n_j - 1) & j \in D \\ \mu_j(n_j) & j \notin D \end{cases}$$

$$\Psi_{jk}(n) = \begin{cases} \alpha_j(n_j - 1) \alpha_k(n_k) & j, k \in D \\ 1 & \text{otherwise.} \end{cases}$$

An easy check shows that Ψ_{jk} are symmetric over any routing-balance partitions $\mathcal{B}_1, \dots, \mathcal{B}_m$ and the ϕ_j 's trivially satisfy (24). Thus, from Theorems 4.6 and 5.1, the equilibrium distribution of X is

$$\pi(n) = c \prod_j w_j^{n_j} \prod_{r=1}^{n_j} \mu_j(r)^{-1} \prod_{\ell \in D} \prod_{r=1}^{n_\ell} \alpha_\ell(r-1) / \eta_\ell(r).$$

The following special case was introduced in [24], and discussed further in [5], [6], [7] and [18]. Suppose that D represents a star-like subnetwork such that node 1 is at the center, $D_0 = D \setminus \{1\}$ and

$$\lambda_{1k} = \lambda_{k1} = \lambda, \quad k \in D_0.$$

Suppose the transition rates of X are

$$q(n, T_{jk}n) = \begin{cases} \lambda \mu_1(n_1) (M_k - n_k) / (a + b n_1) & j = 1, k \in D_0 \\ \lambda_{jk} \mu_j(n_j) & \text{otherwise,} \end{cases}$$

where a, b are constants. This is a particular case of the preceding with

$$\alpha_1(n_1) = 1 \quad \eta_1(n_1) = 1 / (a + b n_1)$$

$$\alpha_k(n_k) = (M_k - n_k) \quad \eta_k(n_k) = 1 \quad k \in D_0. \square$$

6 Palm Probabilities of a Network at its Transitions

A remarkable property of a stationary closed Jackson network process is that whenever a unit moves from j to k , the probability distribution of the disposition of the $N - 1$ unmoved units is the same as the equilibrium distribution π_{N-1} of an identical network with $N - 1$ units in it. We call this the MUSTA property at j, k : “a moving unit sees a time average”. This is similar to the ASTA property (arrivals see time averages) in a queueing system; see [4], [20] and [30]. These transition probabilities concerning the unmoved units are actually Palm probabilities of the process at its transitions. In this section, we present expressions for Palm properties of general events that may occur in a canonical network whenever a transition takes place. We use these to establish several MUSTA properties for general transitions, including a new one for Jackson networks.

We begin with some comments about transitions of a network. A transition or jump of a network process X is synonymous with the movement of a single unit in the network. Each “type” of transition is associated with a subset C of $S \times S$. By a C -transition of X we mean that X jumps from state n to state n' for some $(n, n') \in C$. We find it convenient to express C as

$$C = \{(n + e_j, n + e_k) : (j, k) \in C_n, n \in S_{N-1}\} \quad (29)$$

where $C_n = \{(j, k) : (n + e_j, n + e_k) \in C\}$. In this guise, one thinks of the C -transition as the movement of a unit from node j to node k and the state of the $N - 1$ unmoved units is n , for some $(j, k) \in C_n$ and $n \in S_{N-1}$. Note that C is defined by the sets $C_n, n \in S_{N-1}$, of node pairs for possible movements. We say that C is *state-independent* if C_n is independent of the state n of the $N - 1$ unmoved units. Here are some examples of transitions and their defining sets C_n .

- (a) “A unit moves from a node in the set J to a node in the set K ”:
 $C_n = J \times K$.
- (b) “A jump of X occurs”: $C_n = \{1, \dots, m\}^2$ (this and the preceding transition are state-independent).

- (c) "A unit moves from a node where the most units currently reside":
 $C_n = \{(j, k) : n_j + 1 \geq n_\ell, \ell \neq j; \text{ and } k \neq j\}.$
- (d) "A unit moves into a node that contains less than ν units": $C_n = \{(j, k) : n_k < \nu; \text{ and } k \neq j\}.$
- (e) "The state of the network is in the subset S^* of S and a unit moves into a node with less units": $C_n = \{(j, k) : n_k < n_j \text{ and } n + e_j \in S^*\}.$

For the rest of this section, we assume that X is the network process in Theorem 4.6. We know that the number of occurrences per unit time of a C-transition is given by

$$\begin{aligned}\alpha_C &\equiv \lim_{t \rightarrow \infty} t^{-1} \sum_{s \leq t} 1((X_{s-}, X_s) \in C) \\ &= \sum_{(n, n') \in C} \pi(n) q(n, n').\end{aligned}\tag{30}$$

This can be expressed as follows. Here $\pi_{N-1}(n) = c_{N-1} \Phi(n) \prod w_j^{n_j}$, $n \in S_{N-1}$, is the equilibrium distribution of the network process X with $N-1$ units. **Proposition 6.1** For the network process X in Theorem 4.6, the occurrence rate of a C-transition is

$$\alpha_C = c/c_{N-1} \sum_{n \in S_{N-1}} \pi_{N-1}(n) H_C(n)\tag{31}$$

where

$$\begin{aligned}H_C(n) &= h(n) \sum_{(j, k) \in C_n} w_j \lambda_{jk} \Psi_{jk}(n + e_j) \\ h(n) &= \Phi(n + e_{j_0}) \phi_{j_0}(n + e_{j_0}) / \Phi(n).\end{aligned}$$

Proof. In light of (29) and the form of π and q , expression (30) can be written as

$$\begin{aligned}\alpha_C &= \sum_{n \in S_{N-1}} \sum_{(j, k) \in C_n} \pi(n + e_j) q(n + e_j, n + e_k) \\ &= \sum_{n \in S_{N-1}} \sum_{(j, k) \in C_n} c \Phi(n + e_j) \phi_j(n + e_j) w_j \lambda_{jk} \prod_{\ell} w_\ell^{n_\ell} \Psi_{jk}(n).\end{aligned}$$

From Lemma 4.7(c), we know that

$$\Phi(n + e_j) \phi(n + e_j) = \Phi(n + e_{j_0}) \phi_{j_0}(n + e_{j_0}) = h(n) \Phi(n).$$

Using this and the definition of π_{N-1} in the preceding expression yields (31). \square

We are now ready to describe Palm probabilities of the network process X . General expositions on Palm probabilities appear in [1], [9] and [15]. Assume that X is stationary: that is, $P\{X_t = n\} = \pi(n)$. The *Palm probability of the event $(X_{t-}, X_t) \in A$ given that a C -transition occurs at time t* is defined by

$$P_C(A) \equiv \lim_{u \downarrow 0} P\{(X_{t-u}, X_t) \in A \mid (X_{t-u}, X_t) \in C\}, \quad A \subset C. \quad (32)$$

This probability does not depend on t since X is stationary. The standard definition of a Palm distribution reduces to this simple limit since X is a pure jump Markov process. Here are some special cases of interest:

At a C -transition, the probability that the $N-1$ unmoved units are in state $n \in S_{N-1}$ is

$$P_C(n) \equiv P(\{(n + e_j, n + e_k) : (j, k) \in C_n\}).$$

Whenever a unit moves from node j to node k , the probability that the unmoved units are in state $n \in S_{N-1}$ is

$$P_{jk}(n) \equiv P(\{(n + e_j, n + e_k)\}).$$

Expressions for these Palm probabilities are as follows.

Theorem 6.2 *For the stationary network process in Theorem 4.6,*

$$\begin{aligned} P_C(A) &= \alpha_A / \alpha_C \\ &= \sum_{n \in S_{N-1}} \pi_{N-1}(n) H_A(n) / \sum_{n' \in S_{N-1}} \pi_{N-1}(n') H_C(n'). \end{aligned} \quad (33)$$

In particular,

$$P_C(n) = \pi_{N-1}(n) H_C(n) / \sum_{n' \in S_{N-1}} \pi_{N-1}(n') H_C(n') \quad (34)$$

$$P_{jk}(n) = \pi_{N-1}(n) \Psi_{jk}(n + e_j) / \sum_{n' \in S_{N-1}} \pi_{N-1}(n') \Psi_{jk}(n' + e_j). \quad (35)$$

Proof. From the definition (32) and $A \subset C$, it follows that

$$\begin{aligned} P_C(A) &= \lim_{u \downarrow 0} P\{(X_{t-u}, X_t) \in A\} / P\{(X_{t-u}, X_t) \in C\} \\ &= \lim_{u \downarrow 0} u^{-1} P\{(X_0, X_u) \in A\} / [u^{-1} P\{(X_0, X_u) \in C\}]. \end{aligned}$$

Also, the stationarity of X and the form of $q(n, n')$ imply that

$$\alpha_C = \sum_{(n, n') \in C} \pi(n) q(n, n') = \lim_{u \downarrow 0} u^{-1} P\{(X_0, X_u) \in C\}.$$

The α_A has a similar expression. Putting the preceding expressions together yields (33). Expressions (34) and (35) are just special cases of (33). \square

The Palm probability $P_C(n)$ is the probability that a unit triggering a C-transition “sees” the rest of the units in the state $n \in S_{N-1}$. We will show that in some cases $P_C(n) = \pi_{N-1}(n)$, for each $n \in S_{N-1}$. One can interpret this as a *moving unit sees a time average (MUSTA) at the transition C*; that is, the distribution of the $N - 1$ unmoved units is the same as the distribution of an identical network with $N - 1$ units in it. Similarly, the property $P_{jk}(n) = \pi_{N-1}(n)$ is called *MUSTA at j, k* . In some networks one may have $P_C(n) = \pi_{N-1}(n)$ for some but not all n . We will not consider these partial MUSTA situations.

The next result addresses the question of when does a network have MUSTA properties?

Theorem 6.3 *Suppose the network process X is stationary.*

- (a) *A necessary and sufficient condition for MUSTA at C is that $H_C(n)$ is independent of n .*
- (b) *A necessary and sufficient condition for MUSTA at j, k is that $h(n)\Psi_{jk}(n + e_j)$ is independent of n .*
- (c) *If X has transition rates $q(n, T_{jk}n) = \lambda_{jk}\Phi(n - e_j)/\Phi(n)$, then it has the MUSTA property at each state-independent transition.*

Remark 6.4 Recall that the Jackson network process has the MUSTA property at each j, k . From Theorem 6.3(c), we now know that the Jackson process has the MUSTA property at each state-independent transition.

Proof. (a) If $H_C(n)$ is independent of n , then using this in (34) yields MUSTA at C . Conversely, MUSTA at C and (32) imply that

$$H_C(n) = \sum_{n' \in S_{N-1}} \pi_{N-1}(n') H_C(n'),$$

which is independent of n .

(b) An argument like the preceding one yields statement (b).

(c) Any state-independent transition C can be expressed as

$$C = \{(n + e_j, n + e_k) : (j, k) \in I, n \in S_{N-1}\}$$

for some subset I of node pairs. For such a transition, we clearly have

$$H_C(n) = \sum_{(j,k) \in I} w_j \lambda_{jk}$$

which is independent of n . Thus, by statement (a), the process has MUSTA at C . \square

7 Open Network Processes

The discussion up to now has been on closed networks. In this section, we shall consider an open network in which units enter from the outside and move among the nodes as in a closed network until they eventually exit the network never to return.

Let $X = \{X_t : t \geq 0\}$ denote such a network process with states $n = (n_1, \dots, n_m)$ in $S = \{0, 1, \dots\}^m$ denoting the numbers of units at the m nodes of the network. Let node 0 represent the “outside” of the network. Define $T_{0k}n = n + e_k$ and $T_{j0}n = n - e_j$ or n according as n_j is ≥ 1 or $= 0$. Also, define $T_{jk}n$ as before for $j, k = 1, \dots, m$. We assume that X is a Markov process with transition rates

$$q(n, T_{jk}n) = \lambda_{jk} \phi_j(n) \Psi_{jk}(n), \quad j \neq k \text{ in } \{0, 1, \dots, m\}, \quad (36)$$

where ϕ_j and Ψ_{jk} are functions from S to R^+ , and $\lambda_{jk} \geq 0$ with $\lambda_{jj} = 0$ for each j . Assume that X is irreducible. One can show that this is equivalent to the irreducibility of the Markov routing matrix $\lambda_{jk} / \sum_{\ell=0}^m \lambda_{j\ell}$, $j, k = 0, \dots, m$. Consequently, there exist unique positive numbers w_0, \dots, w_m that satisfy the balance equations

$$\sum_{k=0}^m (w_j \lambda_{jk} - w_k \lambda_{kj}) = 0, \quad j = 0, \dots, m,$$

and $w_0 = 1$.

From the preceding description, it is obvious that this open network process is like a closed network process with one more node 0 with the

special property that $n \pm e_0 = n$. The following is a formal statement to this effect.

Major Remark 7.1 All of the results in sections 4–6 extend *mutatis mutandis* to the open network process. Just replace the closed network notation in the first column below by the notation in the second column.

<i>Closed Network</i>	<i>Open Network</i>
$j, k \text{ in } \{1, \dots, m\}$	$j, k \text{ in } \{0, 1, \dots, m\}$
$S = S_N$	$S = \{0, 1, \dots\}^m$
B_1, \dots, B_m	B_0, \dots, B_m
$n \pm e_j$	$n \pm e_j = 0 \text{ when } j = 0$
$M = N - n_{j_0} \text{ in (24)}$	$M = \sum_{j \neq j_0} n_j$
$c/c_{N-1} \pi_{N-1}(n) \text{ in (31)}$	$\pi(n)$
$\pi_{N-1}(n) \text{ in Section 6}$	$\pi(n)$

Further amplifications on the preceding remark are in order. Although Theorem 4.6 holds as stated, it simplifies considerably as follows for the special case in which $j_0 = 0$.

Theorem 7.2 Suppose the open network process X has transition rates (36) where $\Psi_{j,k}$ are symmetric over routing balance partitions B_0, \dots, B_m and, for each n and $j, k \neq 0$,

$$\phi_j(n) \phi_k(n - e_j) = \phi_k(n) \phi_j(n - e_k).$$

Then the equilibrium distribution of X is $\pi(n) = c \Phi(n) \prod_j w_j^{n_j}$, and X is balanced over B_0, \dots, B_m , where Φ is given by

$$\Phi(n) = \prod_{j=1}^m \prod_{r=1}^{n_j} \phi_0(n - \sum_{s=1}^{j-1} n_s e_s - (r+1)e_j) / \phi_j(n - \sum_{s=1}^{j-1} n_s e_s - r e_j). \quad (37)$$

Remark 7.3 A natural load-dependent network entry rate is $\phi_0(n) = \gamma(n_1 + \dots + n_m)$, where γ is a positive function. In this case (or when $\phi_0(n) \equiv 1$), expression (37) reduces to

$$\Phi(n) = \prod_{j=1}^m \prod_{r=1}^{n_j} \phi_j(n - \sum_{s=1}^{j-1} n_s e_s - r e_j)^{-1}. \quad (38)$$

One might want to use this entry rate and also impose the constraint that the number of units in the network cannot exceed some value M . To do this, simply restrict the state space to $S^* = \{n \in S : n_1 + \dots + n_m \leq M\}$ by defining

$$q(n, n + e_j) = \lambda_{j0} \gamma(n_1 + \dots + n_m) 1(n_1 + \dots + n_m \leq M).$$

Now, if ϕ_j and Ψ_{jk} satisfy the conditions of Theorem 7.2, then this coupled with Theorem 5.1 yields the equilibrium distribution $\pi(n) = c \Phi(n) \prod_j w_j^{n_j}$, $n \in S^*$, where Φ is given by (38). A related example is in [16]. \square

The results on Palm probabilities in Section 6 readily extend to open networks. We say that the open network process has the MUSTA property at the transition C if the distribution of the unmoved units at a C transition is $\pi(n)$ – the same as the (unconditional) equilibrium distribution of the network. An easy check shows that Theorems 6.2 and 6.3, with π_{N-1} replaced by π , are true for the open network.

A common property of a stationary open Jackson network process is that the flows of units exiting the network from the nodes are independent Poisson processes. This result for our network process is as follows. Consider the point process

$$N_j(t) = \sum_{s \leq t} 1(X_s = X_{s-} - e_j) \quad t \geq 0,$$

that represents the number of units that exit the network from node j in the time interval $(0, t]$. Of course, $N_j = 0$ when $\lambda_{j0} = 0$.

Theorem 7.4 *Suppose the open network process X is stationary and has transition rates*

$$q(n, T_{jk}n) = \lambda_{jk} \Phi(n - e_j) \Phi(n)^{-1} \Psi_{jk}(n),$$

where Ψ_{jk} are symmetric over B_0, \dots, B_m and $\Psi_{j0}(n) = 1$ for each j and n . Then N_1, \dots, N_m are independent Poisson processes with rates $w_1 \lambda_{10}, \dots, w_m \lambda_{m0}$, respectively. Furthermore, $\{N_1(s), \dots, N_m(s) : s \geq t\}$ is independent of $\{X_u : u \leq t\}$, for each $t \leq 0$.

Proof. Consider the function

$$\alpha^*(n, z) = \pi(n)^{-1} \sum_{n'} \pi(n') q(n', n) 1(n = T_{j0}n', n' - n = z), \quad n, z \in S.$$

This $\alpha^*(n, z)$ with $z = e_j$ is the intensity of the reversed-time version of N_j ; see [23]. The assertions will follow directly by Theorem 3.2(ii) of [23] if

$$\alpha^*(n, z) = w_j \lambda_{j0} 1(z = e_j).$$

But this is true since

$$\alpha^*(n, z) = \pi(n)^{-1} \pi(n + e_j) q(n + e_j, n) 1(z = e_j),$$

where $\pi(n) = c \Phi(n) \prod_j w_j^{n_j}$ and $q(n + e_j, n) = \lambda_{j0} \Phi(n) / \Phi(n + e_j)$. \square

Some of the flows inside the network may also be independent Poisson processes, see Theorem 4.2 [23], which generalizes the classical results in [3], [19] and [27].

8 Acknowledgement

The author would like to thank K. Edward Chin for his comments on this manuscript and for Example 9.1 in the Appendix.

References

- [1] Baccelli, F. and P. Brémaud(1987), "Palm Probabilities and Stationary Queues," Lecture Notes in Statistics #41. Springer Verlag.
- [2] Basket, E., Chandy, K., Muntz, R. and P. Palacios(1975), "Open, Closed and Mixed Networks with Different Classes of Customers," *J. ACM* 22, 248-260.
- [3] Brémaud, P.(1981), *Point Processes and Queues*, Springer Verlag.
- [4] Brémaud, P.(1988). "Characteristics of Queueing Systems Observed at Events and the Connection Between Stochastic Intensity and Palm Probability", Technical Report.
- [5] Buzacott, J. A. and D. D. Yao(1986), "On Queueing Network Models of Flexible Manufacturing Systems," *Queueing Systems 1*, 29-66.
- [6] Chin, K. E.(1986), "Networks of Queues with State-Dependent Flows," Ph.D. Dissertation, Georgia Institute of Technology.

- [7] Chin, K. E. and R. F. Serfozo(1988), "Networks of Queues with Blocking and Load Balancing," To appear in Proceedings of the Material Handling Research Forum.
- [8] Dallery, Y.(1986), "A Queueing Network Model of Flexible Manufacturing Systems Consisting of Cells", Proceedings of 1986 IEEE International Conference on Robotics and Automation.
- [9] Franken, P., D. König, V. Arndt, and V. Schmidt(1982), *Queues and Point Processes*, Wiley.
- [10] Gelenbe, E. and G. Pujolle(1987), *Introduction to Queueing Networks*, Wiley.
- [11] Gordon, W. J. and G. F. Newell(1967), "Cyclic Queueing Systems with Restricted Queue Lengths," *Operations Research* 15, 266-278.
- [12] Hordijk, A. and N. Van Dijk(1983a), "Adjoint Processes, Job Local Balance and Insensitivity for Stochastic Networks," Proceedings of 44th Session of International Statistics Institute.
- [13] Hordijk, A. and N. Van Dijk(1983b), "Networks of Queues, Part I: Job-local-balance and The Adjoint Process. Part II: General Routing and Service Characteristics," Proceedings International Seminar on Modelling and Performance Evaluation Methodology, INRIA Vol. I, 79-135.
- [14] Jackson, J. R.(1957), "Networks of Waiting Lines," *Operations Research* 5, 518-521.
- [15] Kallenberg, O.(1983), *Random Measures* 3rd Ed. Academic Press.
- [16] Kelly, F. P.(1979), *Reversibility and Stochastic Networks*, Wiley.
- [17] Kingman, J. F. C.(1969), "Markov Population Processes," *J. Appl. Prob.* 6, 1-18.
- [18] Krzesinski, A. E.(1987), "Multiclass Queueing Networks with State-Dependent Routing," *Performance Evaluation* 7, 125-143.

- [19] Melamed, B.(1979), "On Poisson Traffic Processes in Discrete-State Markovian Systems with Applications to Queueing Theory," *Adv. Appl. Prob. 11*, 218-239.
- [20] Melamed, B. and W. Whitt(1988a), "On Arrivals that See Time Averages", Technical report, AT&T Bell Laboratories.
- [21] Melamed, B. and W. Whitt(1988b), "On Arrivals that See Time Averages: A Martingale Approach." Technical Report, AT&T Bell Laboratories.
- [22] Pollett, P.K.(1987), "Preserving Partial Balance in Continuous-Time Markov Chains", *Adv. Applied Probability, 19*, 431-453.
- [23] Serfozo, R. F.(1989), "Poisson Functionals of Markov Processes and Queueing Networks," to appear in *Adv. Appl. Probability*.
- [24] Towsley, D. F.(1980), "Queueing Networks with State-Dependent Routing," *J. ACM 27*, 323-337.
- [25] Van Dijk, N. M. and I. Akyildiz(1988), "Networks with Mixed Processor Sharing Parallel Queues and Common Pools," Technical Report, Georgia Institute of Technology.
- [26] Walrand, J.(1988), *An Introduction to Queueing Networks*, Prentice Hall.
- [27] Walrand, J. and P. Varaiya(1981), "Flows in Queueing Networks: A Martingale Approach," *Math. Operations Research 6*, 387-404.
- [28] Whittle, P.(1968), "Equilibrium Distributions for an Open Migration Process," *J. Appl. Prob. 5*, 567-57.
- [29] Whittle, P.(1986), *Systems in Stochastic Equilibrium*, Wiley.
- [30] Wolff, R. W.(1982), "Poisson Arrivals See Time Averages", *Operations Research 30*, 223-231.

9 Appendix

The following example shows that there does not always exist a unique finest routing-balance partition. The rest of the appendix consists of a few comments on occurrence rates of events in a network.

Example 9.1

Consider the routing intensity matrix

$$\{\lambda_{jk}\} = \begin{pmatrix} 0 & .6 & .1 & .1 & .2 \\ 0 & 0 & .15 & .55 & .3 \\ 0 & .2 & 0 & .1 & .7 \\ .28 & .6 & .08 & 0 & .04 \\ .4 & .2 & .1 & .3 & 0 \end{pmatrix}.$$

This is balanced by the vector $w = (.15, .30, .10, .25, .20)$. The balance-partition \mathcal{B}_π could be either $\{\{1, 2\}, \{3, 4\}\}$ or $\{\{1, 3\}, \{2, 4\}\}$. Clearly there is no finer partition \mathcal{B}_π from which the preceding two partitions can be constructed as unions of sets from it. \square

Considerable information about a network can be obtained by the rates of occurrence of events at its transitions or by the rates of various flows of units among the nodes. The following discussion explains how these rates can be interpreted as limiting averages or as expectations. First, consider the basic event that the network process X is in state n and a unit moves from node j to node k . The number of occurrences of this event per unit time is given by

$$\begin{aligned} r(n, T_{jk}n) &\equiv \lim_{t \rightarrow \infty} t^{-1} \sum_{s \leq t} 1(X_{s-} = n, X_s = T_{jk}n) \\ &= \pi(n)q(n, T_{jk}n). \end{aligned}$$

This is a standard strong law of large numbers for Markov process in which the limit holds with probability one. The sum is over all real $s \leq t$, but only a finite number of the terms being summed are nonzero since X takes only a finite number of jumps in a finite time period. The rate $r(n, T_{jk}n)$ has two other interpretations. Namely, when X is stationary, then

$$r(n, T_{jk}n) = \lim_{h \downarrow 0} h^{-1} P\{X_t = n, X_{t+h} = T_{jk}n\}, \text{ and} \quad (39)$$

$$r(n, T_{jk}n) = E\left[\sum_{t \leq 1} 1(X_{t-} = n, X_t = T_{jk}n)\right]. \quad (40)$$

The latter is the expected number of occurrences of a transition from n to $T_{jk}n$ in a unit interval.

Rates of more complex events at transitions of X can be represented as sums of the basic rates $r(n, T_{jk}n)$. An important example is the *throughput from node j to node k* , or the rate at which units flow from j to k , which is

$$\begin{aligned} r_{jk} &\equiv \lim_{t \rightarrow 0} t^{-1} \sum_{s \leq t} \sum_n 1(X_{s-} = n, X_s = T_{jk}n) \\ &= \sum_n r(n, T_{jk}n). \end{aligned}$$

Similarly, the *throughput of node j* , or rate at which units flow through j , is

$$r_j \equiv \sum_k r_{jk} = \sum_k r_{kj}.$$

These throughputs have two alternate interpretations as in (39), (40).

As another example, suppose that J and K are subsets of nodes. Consider the event that J contains more units than K and a unit moves from J to K . The occurrence rate of this event is

$$\begin{aligned} &\lim_{t \rightarrow \infty} t^{-1} \sum_{s \leq t} \sum_{j \in J} \sum_{k \in K} \sum_{n \in B} 1(X_{s-} = n, X_s = T_{jk}n) \\ &= \sum_{j \in J} \sum_{k \in K} \sum_{n \in B} r(n, T_{jk}n), \end{aligned}$$

where $B = \{n : \sum_{j \in J} n_j > \sum_{k \in K} n_k\}$.



GEORGIA TECH 1885-1985

DESIGNING TOMORROW TODAY

Georgia Institute of Technology
School of Industrial and Systems Engineering
Atlanta, Georgia 30332-0205
(404) 894-2300

June 23, 1989

Dr. Etyan Barouch
AFOSR
Building 410
Bolling Air Force Base, DC 20332-5260

Dear Dr. Barouch:

Enclosed is the final report on my last one-year grant with AFOSR. I am sending you a Ph.D. dissertation under separate cover that was a product of our research program.

Sincerely,

Richard F. Serfozo
Professor

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS			
2a. SECURITY CLASSIFICATION AUTHORITY NA			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.			
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE NA						
4. PERFORMING ORGANIZATION REPORT NUMBER(S)			5. MONITORING ORGANIZATION REPORT NUMBER(S)			
6a. NAME OF PERFORMING ORGANIZATION Georgia Institute of Technology		6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NM			
6c. ADDRESS (City, State, and ZIP Code) ISyE 765 Ferst Dr. NW Atlanta, GA 30332-0205			7b. ADDRESS (City, State, and ZIP Code) AFOSR/NM Bldg 410 Bolling AFB DC 20332-6448			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR		8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER			
8c. ADDRESS (City, State, and ZIP Code) AFOSR/NM Bldg 410 Bolling AFB DC 20332-6448			10. SOURCE OF FUNDING NUMBERS			
			PROGRAM ELEMENT NO. 61102F	PROJECT NO. 2304	TASK NO.	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) Stochastic Flows in Networks						
12. PERSONAL AUTHOR(S) Richard F. Serfozo						
13a. TYPE OF REPORT Final		13b. TIME COVERED FROM 2/16/88 TO 4/15/89		14. DATE OF REPORT (Year, Month, Day) 89,6,15		15. PAGE COUNT
16. SUPPLEMENTARY NOTATION						
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)			
FIELD	GROUP	SUB-GROUP				
19. ABSTRACT (Continue on reverse if necessary and identify by block number) This report summarizes research accomplishments on stochastic flows in networks. The highlight is a new family of probability distributions for describing the numbers of units at the nodes in partially balanced stochastic networks. Such a distribution is a key tool for evaluating the performance and design of a network. Another major accomplishment is the solution of a long-standing problem of finding an expression for the mean time for one unit to move from one sector of a network to another sector. We also developed several models for concurrent movement of units in networks and batch processing at nodes.						
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION			
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Etyan Barouch			22b. TELEPHONE (Include Area Code) (404) 894-2305		22c. OFFICE SYMBOL NM	

Annual Technical Report

Project Title: Stochastic Flows in Networks

Time Covered: One year (February 16, 1988 to April 15, 1989)

Grant No: AFOSR 84-0367

Principal Director: Richard F. Serfozo, Professor
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0205

June 1989

Stochastic Flows in Networks

1 Introduction

This report summarizes our research accomplishments during the last year. The theme of our research is the equilibrium or ergodic behavior of stochastic network processes. Such a process describes the movement of discrete units (customers, parts, data packets etc.) in a network of nodes that process the units. Such processes are often called queueing networks. Some archetypal stochastic networks are as follows:

Computer Networks: Transactions, data packets or programs move among processors, computers, peripheral equipment or files.

Flexible Manufacturing Networks: Parts, tools or material move among a group of work stations and storage areas that machine and store the units for later use or for shipping.

Telecommunications Networks: Telephone calls, data packets or messages move among operators or switching stations.

Maintenance and Logistic Networks: Repairable parts or equipment needed for the operation of a large system move among locations where they are used, repaired, and stored.

Distribution Networks: Goods, orders or trucks move among plants, warehouses or market locations.

Biological Networks: Animals, cells, molecules, neurons, etc. move among locations, states or shapes.

The major concern associated with a stochastic network is to describe the probabilistic behavior of the network in terms of the equilibrium or stationary probability distribution of the numbers of units at the nodes. This distribution is used to derive various performance measures of the network such as the expected cost of operating the network or the percentage of time a sector of the network is overloaded. It is also a basic ingredient for the development of mathematical programming algorithms to select optimal network designs and operating rules. Other important network features include the rate of flow of units on the arcs and through the nodes (the throughputs), the time a unit spends in the network, and the time it takes

for a unit to move from one sector of the network to another.

The existing theory of network processes is primarily for Jackson network processes and its relatives. The key features of these processes are:

- The units move one at a time.
- The nodes operate independently.
- The transition rates depend on only local information: the service rate at a node depends on only the number of units at that node and is independent of the rest of the network.
- The routes of units are independent of each other and hence independent of the congestion in the network.
- The equilibrium distribution is of product form.

By its very nature, however, a network is a system of interacting nodes in which the operation of a node and the routing of a unit may depend on what is happening throughout the network. Examples of dependencies are:

- Parallel or synchronous processing at units of several nodes.
- Alternate routing of units to avoid congestion.
- Accelerating or decelerating the processing rate at a node whenever downstream nodes are starved or congested.
- Units are blocked from entering a sector of the network when the sector cannot handle any more units.

Such dependencies are omnipresent in the networks mentioned above. To model networks with dependencies will require a new generation of network processes that will typically have more complex, non-product-form equilibrium distributions. We have made significant progress in this regard. The following sections give an overview of the results we obtained last year.

2 Markovian Network Processes: Congestion-Dependent Routing and Processing

Our work on this subject will appear in an article with the title above in the journal *Queueing Systems Theory and Applications*. This is a relatively new journal that is to be affiliated with the Operations Research Society of America; I am one of its associate editors. The cumulative papers on this research project is in the appendix.

Prior to this paper, there was no theory for stochastic network processes for networks that have been dependent nodes or congestion-dependent routing. A major hindrance was that no one knew what type of multivariate equilibrium distribution would be appropriate for the numbers of units at the nodes in such a network. The existing theory with product form distributions and reversible ideas was no help in this regard. In this paper, we introduce a wide class of Markovian network processes with system-dependent transition rates that represent a variety of interactions between nodes. This class contains the Jackson processes and essentially all of its generalizations developed to date that have closed-form expressions for its equilibrium distribution. We discovered a new family of multivariate distributions for these processes. These distributions appear to be rather universal and may apply to other multivariate stochastic processes as well. We review the few ad hoc network processes with dependent nodes studied to date and show how they fit into our general framework.

The other topics in the paper are as follows:

Blocking of Certain transitions Due to Network Constraints. We show how to model blocking when the network process is not entirely reversible - prior results required reversibility or used approximations.

Palm Probabilities of a Network at its Transitions. A remarkable property of an open Jackson network is that when a unit moves from one node to another, the probability distribution of the unmoved units is the same as that for the entire network process. That is, a "moving unit sees time averages". We show that this MUSTA property is really an application of Palm probabilities, which are essentially conditional probabilities conditioned on events of zero probability. This was apparently not known before - several papers are now starting to appear on this topic. We describe general Palm probabilities for a variety of transitions and give necessary and

sufficient conditions for the MUSTA property.

Poisson Flows in Networks. The flows of units between two nodes or the departures from a network process are sometimes Poisson processes. Knowing that the departure flow from a network is Poisson is useful when the flow enters another system such as a post-processor or inventory system and one wants to describe the behavior of the auxiliary system. We give easily checkable necessary and sufficient conditions for flows to be Poisson and characterize when they might be independent.

In summary, the major contribution in this paper is the discovery of the new multivariate distribution for representing the equilibrium behavior of Markovian networks. This appears to be a major building block in the theory of network processes with dependent nodes and routings.

3 Partially Balanced Markovian Processes

Much of the work on this topic and the topics in the remaining sections are documented in the Ph.D dissertation by Kwangho Kook titled *Equilibrium Behavior of Markovian Network Processes*, June 1989. We will prepare several papers on these topics this summer.

All of the Markovian network processes studied to date that have known closed-form expressions for their equilibrium distribution satisfy a certain "partial balance" property. Our work described above covers a large subclass of these partially balanced processes, but the entire class of partially balanced processes that have not been characterized is vast in comparison. We initially thought that researchers would develop the equilibrium theory for partially balanced networks by identifying and studying isolated subclasses of these processes. A universal equilibrium distribution for all such processes seemed unlikely. We have found, however, that a universal theory is indeed possible and we have developed much of it only recently. The generic problems we are addressing are:

- (a) Find a general form for the equilibrium distribution of partially balanced networks.
- (b) Find necessary and sufficient conditions on the transition rates of the process for it to be partially balanced.

We have essentially solved these problems for Markovian networks in which units move one at a time. We have also begun to solve them for networks with concurrent movements of units- little is known about these networks, which are described later. This unraveling of the mystery of partially balanced networks is a major breakthrough in the understanding of dependencies in networks. We are extremely pleased with this result.

4 Passage Times in Networks

A long-standing problem in stochastic networks, even for Jackson networks, is to find the mean passage time for a unit to move from one sector of a network to another sector. We have solved this problem for very general processes. This also led us to the study of mean passage times for a variety of routes in a network. The difficulty in this topic is that one cannot approach the problem by standard Markovian reasoning. When a unit begins a passage on a route, it is not known whether the unit will complete the route until the unit reaches the end of the route. In other words, the numbers of units undergoing a passage at any time is a function of the future of the process as well as its past. We overcome this difficulty by a subtle labeling device that allows us to look into the future in a certain regenerative sense. Our expressions for mean passage times provide new performance parameters for assessing the quality of a network.

5 Networks with Concurrent Movements

In actual networks, batch processing and splitting and merging of units are more common than not. These are examples of what we call concurrent movement of units. Little is known about networks with concurrent movements. We have begun to study these networks along the same lines as discussed above.

An important type of concurrent movement is the modeling of resource sharing in a network where the processing at a node requires the use of an auxiliary resource, e.g. computer file, machine tools, pallets. The resources can be represented as artificial units and their normal storage areas as artificial nodes. When a usual unit enters a node for processing, the

artificial units also move simultaneously to the node. When the processing is complete, all the units depart simultaneously. This problem area of resume sharing in networks is relevant to many types of real world problems in manufacturing and computer systems.

6 Service Stations with Batch Arrivals and Batch Services

As a first step in developing the theory of networks with concurrent movements, we studied several typical processing rules for one node in isolation. We eventually developed two models for a service center with batch arrivals and batch services. These models are important in their own right as well as in a network context.

The first model is an $M_n^b/M_n^b/1$ system. This is a single-server system in which batches of units arrive according to a Poisson process with rate depending on the number in the system and the batch sizes are i.i.d. and have a geometric distribution. The neumonic M_n^b refers to this type of process. Similarly, the service process, also represented M_n^b , is a batch service system in which batches depart according to a Poisson process with rate depending on the number in the system and the batch sizes are independent truncated geometric variables. We derive the equilibrium distribution of the number of units in this $M_n^b/M_n^b/1$ system.

One can also interpret the $M_n^b/M_n^b/1$ system as a generalized birth and death process where the births and deaths occur in batches or groups. Our results are therefore applicable in settings where traditional birth and death processes have been used. A special case is the classical $M^b/M/1$ queueing system (M^b means compound Poisson with geometric batches). Other special cases of the $M_n^b/M_n^b/1$ system, which have not been studied before, are the systems $M_n^b/M/s$, $M_n^b/M/\infty$, $M_n^b/M_n/1$, $M_n/M_n^b/1$, etc. (M_n means state-dependent Poisson Process).

The second model we study in this chapter is an $M^b/M^B/1$ system. This is a single server system in which batches of units arrive according to a Poisson process and the batch sizes are i.i.d. geometric variables. The service process is a batch service system in which the B units are served together, except when less than B units are in the system and ready for

service, at which time all units are served. The service time for a batch is exponentially distributed. We derive the equilibrium distribution of the number of units in this system. In our future research, networks with concurrent movements where the nodes operate like these batch service systems.

7 Appendix

Cumulative List of Papers During Grant Period 9/84 to 4/89.

All but the last one are authored by R. F. Serfozo.

Partitions of point processes: multivariate Poisson approximations, *Stoch. Processes Appl.* 20, 281-294, 1985.

Compound Poisson approximations for sums of random variables, *Ann. Probability* 14, 1391-1398, 1986.

Heredity of stationary and reversible stochastic processes, *Adv. Appl. Probability* 18, 574-576, 1986.

Equitable transit charges in multi-administration telecommunications networks, *Queueing Systems: Theory and Applications* 2, 83-92, 1987.

Point Processes. Technical Report, Center for Stochastic Processes, University of North Carolina, 1987. Part of ORSA handbook on stochastic processes.

Extreme values of birth and death processes and queues, *Stoch. Processes Appl.* 27, 291-306, 1988.

Extreme values of queue lengths in M/G/1 and G/M/1 systems, *Math. Oper. Res.* 13, 349-357, 1988.

Poisson functionals of Markov processes and queueing networks, *Adv. Appl. Probability* (to appear) 1989.

Markovian network processes: congestion-dependent routing and processing, *Queueing Systems* (to appear) 1989.

Networks of queues with blocking and load balancing, joint with K.E. Chin, *Proceedings of Material Handling Forum* (to appear) 1989.

Equilibrium Behavior of Markovian Network Processes, by Kwangho Kook, Ph.D. Dissertation in Industrial and Systems Engineering, Georgia Institute of Technology. 1989.